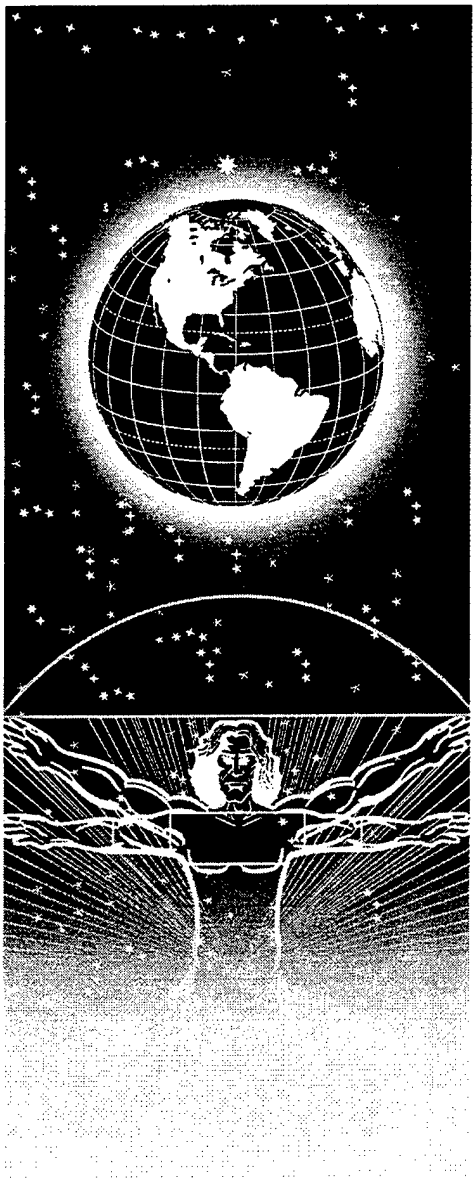
AL/AO-TR-1997-0094



UNITED STATES AIR FORCE ARMSTRONG LABORATORY

THE DEVELOPMENT OF THE SUSTAINED OPERATIONS ASSESSMENT PROFILE (SOAP)

Paul D. Retzlaff
Raymond E. King
Royden W. Marsh
Jonathan French

AEROSPACE MEDICINE DIRECTORATE
CLINICAL SCIENCES DIVISION
NEUROPSYCHIATRY BRANCH
2507 Kennedy Circle
Brooks Air Force Base, TX 78235-5117

July 1997

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this technical report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

ROYDEN W. MARSH, MD
Project Scientist

DANIEL L. VAN SYOC, Lt Col, USAF, MC, CFS
Acting Chief, Clinical Sciences Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1997	Interim Technical Report, Aug 96 - Jul 97

4. TITLE AND SUBTITLE
The Development of the Sustained Operations Assessment Profile (SOAP)

5. FUNDING NUMBERS

6. AUTHOR(S)
Paul D. Retzlaff
Raymond E. King
Royden W. Marsh
Jonathan French

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Armstrong Laboratory
Aerospace Medicine Directorate
Clinical Sciences Division, Neuropsychiatry Branch
2507 Kennedy Circle
Brooks Air Force Base, TX 78235-5117

8. PERFORMING ORGANIZATION REPORT NUMBER
AL/AO-TR-1997-0094

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This work describes the development of a new psychological test to assess fatigue in military and civilian personnel engaged in sustained operations. The Sustained Operations Assessment Profile (SOAP) was developed through the integration of clinical experience, psychometric methods, and empirical analysis. Using a sample of 84 subjects, a thorough test development plan was accomplished. The 10 scales cover three broad areas of functioning including cognitive, affective, and arousal dimensions. The scales have normative data and are demonstrated to be reliable and valid. Further validation work is suggested to improve the test.

14. SUBJECT TERMS
Assessment
Fatigue
Pilots
Psychological testing
Sustained operations

15. NUMBER OF PAGES
36

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

PREFACE

This project was funded through the Air Force Office of Scientific Research (AFOSR) Summer Faculty Research Program.

SUMMARY

This work describes the development of a new psychological test to assess fatigue in military and civilian personnel engaged in sustained operations. The Sustained Operations Assessment Profile (SOAP) was developed through the integration of clinical experience, psychometric methods, and empirical analysis. Using a sample of 84 subjects, a thorough test development plan was accomplished. The 10 scales cover three broad areas of functioning including cognitive, affective, and arousal dimensions. The scales have normative data and are demonstrated to be reliable and valid. Further validation work is suggested to improve the test.

The Development of the
Sustained Operations Assessment Profile (SOAP)

When the enemy sees an advantage but does not
advance on it, they are weary.

Sun Tzu
The Art of War

INTRODUCTION

The great Chinese warrior/philosopher Sun Tzu frequently points out in his book, The Art of War, written some 2,000 years ago, that the outcome of battle is often in favor of those who best manage fatigue, "those who master energy". Fatigue derived from inadequate sleep or from working through the normal sleep period can and has led to great catastrophes, both military and civilian. As today's military and other "down-sized" organizations are pressed to do more with less, they will increasingly be faced with reducing the threat from fatigue degraded performance. Extended on-duty demands is referred to as Sustained Operations (SusOps), during which a long period on duty is required with, typically, no sleep breaks. An example might be ferrying fighter aircraft overseas, often requiring a 12-16 hour flight in a single seat fighter where sleep is impossible. There are parallels in the civilian world too for long duration job requirements such as for emergency workers during floods or fires. Managing fatigue carefully is often vital to the success of these extended duty days. Important corollaries of fatigue management are the ability to recognize and to estimate the fatigue levels of those involved so that the timing of sleep breaks can best be determined. This paper describes the validation of a survey tool, developed for the US Air Force, that can be used in the field or laboratory to rapidly assess fatigue and other related aspects of mental health, performance, and mental status.

The Air Force mission often requires preparation for many long duration activities (Ferrer, Bisson, and French, 1995). For example, the Air Force has closed many of its forward bases overseas and must increasingly rely on long duration flights from the US to meet it's global commitments. Frequently, these operations require multiple days of SusOps with little sleep or poor sleep in between missions (Neville, Bisson, French, Boll, and Storm, 1994). As well, an increased reliance on night operations will compound the problem of inadequate sleep if not managed properly (French, Bisson, Neville, Mitcha, and Storm, 1994). The better crews manage this fatigue the more effective they will be in completing their mission.

The effects of sleep deprivation on reaction time and accuracy in laboratory studies of cognitive ability are well known (for example, Dinges and Kribbs, 1991). Compounding these degradations in performance due to extended workdays are circadian effects on cognition (Tilley and Brown, 1992). Dramatic deficiencies in operational measures are also easily observed in simulator tests, for example in high fidelity aircraft simulators (Moore-Ede, 1993). These studies report that long duration missions and

circadian factors seriously degrade behaviors, like flying, that are often automatic and well established. The ability to respond to changing conditions is also impaired. Microsleeps, or involuntary naps, even of brief duration, are especially dangerous at the speeds of today's aircraft and other complex systems.

While brief periods of fatigue and poor sleep hygiene result in decrements in cognition and affect, some missions are always "up tempo". NASA shuttle missions, for example, combine one to two weeks of high operational tempo with acutely fatiguing, novel situations. The typical astronaut trains for two years for a mission. This training is scheduled down to the hour and becomes almost onerous in the months just prior to a launch. The excitement combined with last minute changes and launch postponements can be very fatiguing. As this report goes to press, Mir, the Russian space station, has just incurred a damaging on-orbit collision. The astronauts are in the position of being faced with life-threatening danger on the heals of a fairly chronic mild fatigue from the close confines, the usual necessary repairs, and the demands of a unique, hostile environment. The air conditioning, lights, and toilet are off. In this most critical of periods, the astronauts are probably at their worst; fatigue may easily lead to mistakes and tragedy. Indeed, one wonders the degree to which fatigue may have contributed to the collision itself.

Traditional means to study fatigue have been objective performance tests and standardized mood surveys like the Stanford Sleepiness Scale (SSS; Hoddes, Zarcone, Smythe, Phillips, and Dement, 1975), the Profile of Mood Survey (POMS; McNair, Lorr, and Droppelman, 1971) or Visual Analog Scales (VAS; Folstein and Luria, 1973). These metrics have distinct advantages and disadvantages. Objective performance tests, while often extremely fatigue sensitive, require stable performance levels to be achieved so that deviations from normal can be discerned (Schnieder, 1985). This can take hours of practice and such demands of crew time are frequently difficult to achieve. The SSS is extremely convenient but consists of only a single question with single response and fails to tap into other mood dimensions. The use of single items also results in limited reliability from a psychometric perspective. The POMS reliably measures 6 mood dimensions (fatigue, vigor, confusion, anger, tension, depression) but consists of 65 adjective evaluations. These take considerable time to administer and to score. VAS are frequently shorter in time requirement and allows the respondent to place a mark anywhere on a line between two extremes to indicate a response. While this allows for a wider range of scores than either of the other mood measures, this can leave the questions and the unguided nature of the responses open to a considerable differences in interpretation. The resolution of this scale may also be overly artificial. The SSS and the VAS are typically single responses for each mood. The POMS has several questions to assess each mood and may thus strengthen the reliability of the mood assessment. A metric that reliably assessed multiple mental health dimensions with regard to fatigue states would be extremely useful for operational field and laboratory tests.

PURPOSE

The purpose of this psychological test is to operationalize and assess dimensions of interest with regard to the mental and physical health of military crew engaged in sustained operations. It is intended to quickly (3 minutes) assess a broad range of variables across cognitive, affective, and arousal areas. The reasons to construct a new test are to ensure: 1) a standard assessment format across variables, 2) the high reliability of the scales, and 3) the potential for highly valid variable assessment.

SOAP Description

Format The Sustained Operations Assessment Profile (SOAP) is a 90-item test. Appendix A provides a copy of the test form. The test may be printed on the front and back of a legal size piece of paper. As such, it is very compact and appears "short" to the subject. The subject is requested to endorse the items on a 1 to 5 scale. Anchors are "Not at all" to "Very much".

Each of the 10 scales have 9 items which represent related aspects of the dimension. The items are presented as blocks under each of the scale names. This format differs from the typical psychological test which usually randomizes items across scales and does not make explicit the scale names.

The test usually only takes between 3 and 5 minutes to complete. It is therefore well suited to situations requiring repeated administrations.

Scales The scales of the test include three cognitive dimensions (Poor Concentration, Boredom, and Slowed Reactions), three affective dimensions (Anxiety, Depression, and Irritability), and four arousal dimensions (Fatigue/ Low Energy, Poor Sleep, Work Frustration, and Physical Discomfort).

Cognitive Dimensions

1. Poor Concentration: High scorers report difficulty concentrating and paying attention. They have difficulty engaging in concentrated effort and consequently work inefficiently. They must repeat work and pay extra attention to complete tasks.

2. Boredom: High scores are disinterested and complacent. They view the work as tedious and tiring. They no longer see "fun" in the workload and time appears to have slowed for them.

3. Slowed Reactions: High scorers are cognitively and physically slowed. Additional effort is required to keep up with the task load. Things around them seem slowed and they are also slowed. Motor activity is impaired by the cognitive inputs and outputs.

Affective Dimensions

4. Anxiety: High scorers feel nervous, anxious, and worried. Physically they are tense and jittery. Autonomically, they are vigilant and upset.

5. Depression: High scorers feel depressed, unhappy, and sad. They are discouraged and lacking in enjoyment. They are concerned about their feelings and beginning to feel helpless.

6. Irritability: High scorers are annoyed with others and feeling unfriendly. This includes the desire to be alone and away from others. They are impatient, disagreeable, and, may be, angry.

Arousal Dimensions

7. Fatigue / Low Energy: High scorers are tired to the point of feeling burned out and worn out. They lack energy and see little ability to go on or perhaps even move.

8. Poor Sleep: High scorers report sleep that was too little and of poor quality. This leads to feeling sleepy and desiring sleep. Their eyes are closing and they may be nodding off. To compensate, they may be focusing on keeping their eyes open and bouncing or shaking.

9. Work Frustration: High scorers are tired of working. They wish they didn't have to complete tasks and hope nothing new comes up. The work feels like a grind and drudgery. They may feel they need help to complete the job.

10. Physical Discomfort: High scorers have sore muscles and are stiff. They want to get up and stretch. They may feel physically uncomfortable due to perspiration and binding clothing/ gear. They feel the discomfort in their arms and legs, head, and eyes.

Construction Plan

This test was developed using the Domain Theory test construction model (Nunnally, 1978). There are three steps in this procedure. The first step is a clinical stage wherein domains (scales) and items are generated based upon the experience of a range of experts. A set of scales is decided upon based on clinical and research experience. Then, a group of items is written by a number of authors for each of the scales.

The second step is a statistical, internal procedure used to ensure the reliability of the test. Here a sample of subjects are given the test. Correlations are calculated between each of the items and the total score for each scale. The general hope is that items written for a scale will each correlate highly with the total sum of all items in the scale. This procedure ensures that each item contributes to a specific dimension. Usually a number of items are eliminated at this point because they are empirically inconsistent with the scale. The final set of items are then subjected to internal consistency analysis. Here a Cronbach Alpha is calculated for each scale as the metric of reliability.

The third and final stage involves validity and uses external statistical techniques. Here the scales may be given to two samples of known difference and it is hoped that the test will differentiate the two groups. This would be a criterion validity approach. Additionally, the scales may be correlated with other tests of similar type. This would be a form of construct validity.

The purpose of these three steps is to ensure that the test maps important domains, is reliable, and is valid. The three steps in this particular order have the tendency to extract from each step the valuable elements and compensating for the weaknesses of each of the three individual steps.

Subjects

Three groups of subjects were used to develop the SOAP. The first group included 33 Air Force enlisted members and officers. About half were male and ages ranged from 18 to the mid-50's.. They took the SOAP between 1500 and 1600 on the day of an early morning recall. The recall had begun at about 0300 and most reported to base by 0400. As such, this group had been awoken during the middle of the night, had disrupted sleep, and had then worked for about 12 hours.

The second group consisted of 41 college undergraduates and served as a control group. Again about half were male but the ages were somewhat younger with a range of 19 to 34. This group took the test at 0900 during a class meeting time. While at first blush this group may be seen as a well rested control, it is important to remember the lifestyle of the average college student.

The final group of 10 took the test twice. They were a safety investigation board which spent over a month investigating a transport accident in the Rocky Mountain area. The majority of subjects were male and ages were between 26 and the mid-50's. They took the test on the 20th day of the investigation once in the morning at about 0900 and again that night at about 2100. The safety board usually met every day including weekends and usually worked in excess of 12 hours per day. As such, this group is probably both chronically and acutely fatigued.

Scale and Item Development

Domain (scale) development was accomplished through the use of a number of clinical and research experts. A number of experts were interviewed with a eye toward the delineation of a number of domains which have proven important across an array of military and civilian sustained operations experiences. These included those familiar with general Air Force aviation, extended Air Force bomber missions, the NASA shuttle program, submarine missions, and military mobility exercises. An initial list of scales was circulated and experts eliminated, added, and combined scales as they believed the phenomenology of fatigue behaved. A final list of 10 scales was fairly well supported by most experts.

After the scales were delineated, the experts were then used to write items which mapped to each of the domains. Items written by each expert were then reviewed and edited by others. The 9 items viewed as "best" were included in the initial form of the SOAP. This procedure resulted in 10 scales with 9 items for each scale. This number of scales and number of items per scale was guided by the desire to have a final test that tapped a number of dimensions of fatigue but that was also short enough to allow for easy administration.

Item-total correlations (Appendix B) were calculated using a mixed sample of 84 subjects. These correlations indicated that all items on each scale were highly correlated with each scale. It is far more typical in test construction to find a number of items which must be eliminated or re-written. As such, all items were retained for the final form of the test.

The weakest items in the test were two under Physical Discomfort. The first with an item-total correlation of .54 was the endorsement of "skin sticky/ dirty". This relatively low correlation is probably the result of very few of the subjects experiencing this sensation. None of these subjects were in field conditions nor on several day missions. The second item, "what I'm wearing is uncomfortable", is probably low (.46) because none of the subject samples were in unusual garments or gear. Also none had been wearing the same outfit for more than 12-14 hours. While an argument could be made that these items should be changed or removed, they will probably prove valuable in situations where uncomfortable outerwear is more common such as the multiday wearing of flight suits or chemical weapons protection.

Norming

Table 1 presents the means and standard deviations for the four testings of the three groups. Statistical differences will be examined later under the validity sections but it is important to look at the relative magnitude of scores. All scale scores can range from 9 (answering 1 to each of the 9 items) to 45 (answering 5 to all items). The means found in the table are generally in the middle third of the possible range. Therefore, there is no truncation of scores due to "ceiling" or "floor" effects. The standard deviations are in the 8 to 10 point range and suggest a sufficient amount of variance. It is not the case that all subjects answered any of the scales in the same way. These descriptive statistics indicate that the scales behave quite well from a purely metric perspective.

The table also indicates general levels across the four testings. The college student group is more often than not the lowest of the four means. The recall group is highest on 3 scales and the second testing of the board is highest on the rest of the scales. In general this is consistent with the a priori expectation of the behavior of the three subject samples.

Table 2 provides initial norming work. Here the recall, student, and second testing of the board are combined into grand descriptive data. With 41 students and a total of 43 fatigued military subjects, these norms may be viewed as a mixed norm sample with approximately half non-fatigued and half fatigued. One inference could be that the mean for each scale probably does a good job of separating the two groups. An interpretation, therefore, could be that subjects with scores above the mean are fatigued and those below the mean are not fatigued.

Additionally, a "cut score" is provided in this table. Here the frequency distributions of the data were analyzed to find the 95th percentile. The raw score for the 95th percentile was abstracted for these cut scores. Here the analysis goes that anyone scoring above these cut scores is very severely fatigued and probably require administrative or clinical attention.

Further research with the SOAP should allow for better norms and situations specific cut scores.

Reliability

Reliability is the freedom from error that a scale enjoys. In classical test theory it is felt that the most probable source of error is in the selection of items. In essence, while there are several sources of error in a score, the biggest is probably the error associated with poorly conceptualized or worded items. The theory goes that if the items all relate to one specific and homogeneous domain, the test will be relatively free of erroneous variance.

The first step in this is the item-total correlations noted above. The second part looks at the scales as wholes rather than at an item level. Here each scale is subjected to a procedure which quantifies the internal consistency of the scale. The Cronbach Alpha is the coefficient. Typical scales within psychology generally have reliabilities in the .70 to .90 range (higher is better).

Table 3 provides the internal consistencies for the SOAP scales. As can be seen, all scales are in the .90's with the exception of Physical Discomfort at .88. These are extremely high. They indicate that each scale is measuring only one thing and with very little error.

To be fair, there is some artifact driving up these coefficients. As mentioned above, most psychological test tend to mix items up within the test form so that scale items are not arranged together. This has been done to keep subjects from sensing what being assessed. In some situations, such as psychiatric disability evaluation for compensation, doing so is important. The SOAP, however, was constructed in such a manner as to orient the subjects to the dimensions by actually including scale names on the test form and arranging all items within a scale together. The problem with this is that the subjects may form an opinion as to their "Poor Concentration" and then endorse the

items on that general basis as opposed to the specific items. This can and probably did drive the Cronbach Alphas up by artifact. The converse of this argument is that if subjects actually know what an item relates to they may misinterpret individual items less and actually produce less item related error. It is safe to say that the scales of the SOAP are very reliable.

Internal Construct Validity

Table 4 provides an intercorrelation matrix of the SOAP scales. There is a good deal of intercorrelation among the scales of this test. To some degree, that is expected as all scales are elements of fatigue. Then again, if scales are too highly correlated they cease to be individual and specific. Only the recall group and the second testing of the safety board were used here to model the behavior of the test in fatigued samples.

The correlations are highest among scales of similar dimensionality. All 3 of the cognitive dimensions scales are intercorrelated above .60. All 3 affective dimensions scales are correlated above .61. The arousal dimensions are less highly correlated especially the Work Frustration scale. Correlations across dimension type are lower usually less than .51. These patterns are logical and within the normal limits of tests of this type.

Table 5 provides the summary data for a factor analysis of the 10 scales using the above intercorrelation matrix. Three Eigenvalues exceed 1.00 so three factors were extracted and rotated. The resulting factor solution is amazingly close to the initial clinical theory holding that there would be cognitive, affective, and arousal components. All scales of similar dimension highly and positively loaded on the hypothesized dimension with the exception of Work Frustration. Work Frustration loaded about equally on the cognitive factor and the affective factor. As such, Work Frustration should probably be viewed as not an arousal component but as a mixture of the other two.

The factor solution accounted for a remarkable 78% of the covariance. The rotation of the factors very nicely spread it across the three factors. It is also important to note that the test form did not relay to the subjects these three dimensions. Therefore, this validation of the three dimension theory is free of any artifact.

Operational Validity

In order for tests to be considered valid, the scales must not only behave well internally, but the test must actually differentiate among subjects of known status. To that end, the recall group will be compared to the student group and the two testings of the safety board will be compared.

The recall group was significantly higher than the student group (Table 6) on four of the scales. They were significantly higher on Poor Concentration, Boredom, Slowed Reactions, and Poor Sleep. This points to the fact that the scales not only map the early

morning recall in the Poor Sleep scale but also the cognitive impairment of the sample. In essence, you can drag troops in from their beds in the middle of the night but you can't make them think.

The safety board, again, took the test about 12 hours apart before and after a long day of board proceedings. Table 7 shows that this group became significantly poorer at concentration and began to feel more depressed and more irritable. This is probably consistent with most peoples' experience working on committees. In the case of the safety board, it can be seen that long days lead to increasing difficulty concentrating and also a very affective reaction. The Depression and Irritability scores show a level of affect inconsistent with good proceedings.

These two operational validation studies together point of the specificity of fatigue given differing situations. It is not the case that the two fatigued groups simply became worse on all scales. There was a very specific pattern of fatigue. It is hoped that the SOAP will be able to model such specific patterns of fatigue in other samples of know situation.

External Construct Validity

The college sample of 41 subjects also took the ALAPS (Retzlaff, King, McGlohn, and Callister, 1996). The ALAPS is a psychological test optimized for pilots. It includes scales of personality, psychopathology, and crew interaction. Two of the scales are entitled Anxiety and Depression. It is, therefore, possible to validate these two constructs in the SOAP. As can be seen in Table 9, The SOAP Anxiety scale correlates with the ALAPS Anxiety scale at .77. Further, the SOAP Depression scale correlates with the ALAPS Depression scale at .87. Both of these are very high and indicate excellent convergent construct validity. Of interest is the fact that out of the entire 10 (SOAP) by 15 (ALAPS) intercorrelation matrix, these two correlations were the highest two. This points to a degree of divergent construct validity and specificity.

CONCLUSIONS

The need for a psychometrically sound test of fatigue dimensions has been presented. To that end, a new test of fatigue and related mental health variables has been developed. The Sustained Operations Assessment Profile (SOAP) has variables of interest, is reliable, and is valid.

Additional psychometric research on the SOAP is needed. Studies of the stability of its scales are important. While designed to assess relatively unstable and situational dimensions, some evidence of the stability of scores in test-retest envirnments is needed. Additional internal consistency reliability work with larger samples would improve estimates of reliability. Finally, further validity work is indicated. Such validity work should include construct validity analyses against other tests and further operational validity work mapping the behavior of the test in various fatiguing situations.

It is currently sufficiently well developed to recommend its use in laboratory and field experiments. It is also suggested that it is appropriate for a large range of settings including the assessment of aircrew in operational settings, personnel in intense training situations such as survival school, and any other Air Force personnel engaged in fatiguing situations. Other groups who may benefit from it are astronauts, Navy personnel such as SEALS or submarine crews, Army, and Marine groups. The test allows for the objective assessment of fatigue states and as such will benefit those involved in research as well as those engaged in establishing operational requirements where sustained operations are present.

REFERENCES

Dinges, D. F. and Kribbs, N. B. (1991). Performing while sleepy: Effects of experimentally-induced sleepiness In T.H. Monk (ed). *Sleep, sleepiness and performance (pp97-128).* New York, John Wiley.

Ferrer, C.F., Bisson, R.U and French, J. (1995). Circadian rhythm desynchronosis in military deployments: A review of current strategies. *Aviation, Space and Environmental Medicine, 66,* 571-578,

Folstein, M.D. and Luria, R. (1973). Reliability, validity and Clinical application of the visual analogue mood scale. *Psychophysiology, 3,* 479-486.

French, J. Bisson, R.U., Neville, K.J., Mitcha, J. and Storm, W.F. (1994). Crew fatigue during simulated, long duration B-1B bomber missions. *Aviation, Space and Environmental Medicine, 65,* A1-A6

Hoddes, E., Zarcone, V.,Smythe, H., Phillips, R. and Dement, W.C. (1975). Quantification of Sleepiness : A new approach. *Psychophysiology, 10,* 431-436.

McNair D., Lorr, M. and Droppelman, L. (1975). *Edits manual for the Profile of Mood States.* San Diego: Educational and Industrial Testing Service.

Moore-Ede, M. (1993). *The Twenty-Four Hour Society, Chapter 6.* Reading, MA.: Addison Wesley Publishing .

Neville, K.J., Bisson, R.U., French, Boll, P.A. and Storm, W.F. (1994). Subjective fatigue of C-141 aircrews during Operation Desert Storm. *Human Factors, 36(2),* 339-349.

Nunnally, J. C. (1978). *Psychometric Theory.* New York: McGraw-Hill.

Retzlaff, P. D., King, R. E., McGlohn, S. E., & Callister, J. D. (1996). *The development of the Armstrong Laboratory Aviation Personality Survey (ALAPS).* (AL/AO-TR-1996-0108). Brooks AFB: Armstrong Laboratory.

Schnieder, D. (1985). Training high performance skills: Fallacies and guidelines. *Human Factors, 27,* 285-300.

Tilley, A. and Brown, S. (1993). Sleep deprivation. In A.P. Smith, and D.M. Jones (Eds) *Handbook of human performance: Vol 3. State and trait* (pp 237-259). San Diego: Academic Press.

Table 1

Descriptive Statistics for SOAP Development Samples.

Scale	Recall Group Mean (sd)	Student Group Mean (sd)	Board AM Mean (sd)	Board PM Mean (sd)
Cognitive Dimensions				
1. Poor Concentration	22.1 (7.6)	17.7 (8.0)	19.1 (4.0)	25.6 (6.0)
2. Boredom	25.2 (9.2)	18.0 (7.8)	19.5 (5.7)	23.7 (8.2)
3. Slowed Reactions	18.6 (8.1)	14.7 (6.5)	17.6 (4.5)	21.9 (5.5)
Affective Dimensions				
4. Anxiety	16.9 (7.5)	16.9 (8.3)	16.7 (3.7)	19.8 (4.9)
5. Depression	17.3 (7.9)	15.4 (8.3)	17.0 (7.7)	21.7 (8.5)
6. Irritability	18.0 (8.0)	16.2 (8.0)	18.0 (6.4)	21.3 (7.0)
Arousal Dimensions				
7. Fatigue / Low Energy	22.3 (8.9)	18.7 (10.4)	21.6 (5.3)	24.3 (6.6)
8. Poor Sleep	22.5 (8.6)	17.6 (6.9)	21.6 (8.4)	21.4 (7.0)
9. Work Frustration	24.1 (11.5)	20.9 (10.4)	20.3 (8.8)	21.4 (11.9)
10. Physical Discomfort	17.7 (6.6)	15.8 (8.0)	17.3 (7.1)	14.6 (5.3)
N =	33	41	10	10

Table 2

Norms for the SOAP Scales.

	Mean(sd)	Cut Score
Cognitive Dimensions		
1. Poor Concentration	20.3 (8.1)	34
2. Boredom	21.5 (9.0)	36
3. Slowed Reactions	17.1 (7.5)	31
Affective Dimensions		
4. Anxiety	17.2 (7.7)	31
5. Depression	16.9 (8.3)	33
6. Irritability	17.5 (8.0)	30
Arousal Dimensions		
7. Fatigue / Low Energy	20.8 (9.6)	39
8. Poor Sleep	20.0 (7.9)	35
9. Work Frustration	23.1 (10.8)	40
10. Physical Discomfort	17.2 (7.6)	29

Note: N = 84, Cut Score = 95 percentile.

Table 3

Internal Consistencies for the SOAP Scales.

	Alpha
Cognitive Dimensions	
1. Poor Concentration	.94
2. Boredom	.95
3. Slowed Reactions	.93
Affective Dimensions	
4. Anxiety	.92
5. Depression	.93
6. Irritability	.94
Arousal Dimensions	
7. Fatigue / Low Energy	.95
8. Poor Sleep	.90
9. Work Frustration	.96
10. Physical Discomfort	.88

Note: N=84

Table 4

Intercorrelation Matrix for SOAP Scales

	1	2	3	4	5	6	7	8	9
Cognitive Dimensions									
1. Poor Concentration	-								
2. Boredom	66*	-							
3. Slowed Reactions	77*	61*	-						
Affective Dimensions									
4. Anxiety	28-	37-	28-	-					
5. Depression	50-	67*	50-	62*	-				
6. Irritability	42-	51-	37-	68*	68*	-			
Arousal Dimensions									
7. Fatigue / Low Energy	58	59	63*	48-	65*	48-	-		
8. Poor Sleep	48-	44-	55	39-	40-	43-	67*	-	
9. Work Frustration	54	69	40-	48-	63*	55	55	44	-
10. Physical Discomfort	41-	29-	46-	51	46-	48-	72*	56	34

note: N=43, * = correlation >.59 and - = correlation < .51.

Table 5

Factor Analysis of SOAP Scales

	Factor 1	Factor 2	Factor 3
Cognitive Dimensions			
1. Poor Concentration	.83*	.13	.33
2. Boredom	.81*	.43	.08
3. Slowed Reactions	.74*	.04	.50
Affective Dimensions			
4. Anxiety	.00	.83*	.34
5. Depression	.45	.73*	.22
6. Irritability	.22	.80*	.26
Arousal Dimensions			
7. Fatigue / Low Energy	.45	.35	.70*
8. Poor Sleep	.34	.19	.73*
9. Work Frustration	.58	.61	.07
10. Physical Discomfort	.07	.33	.85*
Variance Accounted For	28%	27%	23%

note: N=43, * = Loadings > .69.

Table 6

Group Differences between Recall Group and Student Group.

Scale	Recall Group Mean (sd)	Student Group Mean (sd)	t (df 72)	p
Cognitive Dimensions				
1. Poor Concentration	22.1 (7.6)	17.7 (8.0)	2.41	.0185*
2. Boredom	25.2 (9.2)	18.0 (7.8)	3.63	.0005*
3. Slowed Reactions	18.6 (8.1)	14.7 (6.5)	2.32	.0232*
Affective Dimensions				
4. Anxiety	16.9 (7.5)	16.9 (8.3)	0.05	.9635
5. Depression	17.3 (7.9)	15.4 (8.3)	0.99	.3252
6. Irritability	18.0 (8.0)	16.2 (8.0)	0.98	.3295
Arousal Dimensions				
7. Fatigue / Low Energy	22.3 (8.9)	18.7 (10.4)	1.58	.1190
8. Poor Sleep	22.5 (8.6)	17.6 (6.9)	2.69	.0089*
9. Work Frustration	24.1 (11.5)	20.9 (10.4)	1.25	.2140
10. Physical Discomfort	17.7 (6.6)	15.8 (8.0)	1.09	.2809
N=	33	41		

note: * = $p < .05$.

Table 7

Differences across Day for Board Group.

Scale	Board AM Mean (sd)	Board PM Mean (sd)	t(df9)	p
Cognitive Dimensions				
1. Poor Concentration	19.1 (4.0)	25.6 (6.0)	2.65	.0266*
2. Boredom	19.5 (5.7)	23.7 (8.2)	2.22	.0534
3. Slowed Reactions	17.6 (4.5)	21.9 (5.5)	1.80	.1055
Affective Dimensions				
4. Anxiety	16.7 (3.7)	19.8 (4.9)	1.66	.1318
5. Depression	17.0 (7.7)	21.7 (8.5)	3.69	.0050*
6. Irritability	18.0 (6.4)	21.3 (7.0)	3.41	.0077*
Arousal Dimensions				
7. Fatigue / Low Energy	21.6 (5.3)	24.3 (6.6)	1.41	.1920
8. Poor Sleep	21.6 (8.4)	21.4 (7.0)	-.10	.9191
9. Work Frustration	20.3 (8.8)	21.4 (11.9)	0.21	.8371
10. Physical Discomfort	17.3 (7.1)	14.6 (5.3)	-.86	.4129

note: N=10, * = <.05.

Table 8

SOAP Correlations with ALAPS

SOAP Scale	ALAPS Scale	Correlation
Anxiety	Anxiety	.77
Depression	Depression	.87

Note: N=41.

Appendix A

SOAP Form

SOAP (VER. 96.6)

Please rate each of the following as they apply to you **for about the last hour or so**, 1 means not at all, while 5 means very much.

1. POOR CONCENTRATION

	NOT AT ALL				VERY MUCH
1. DIFFICULTY CONCENTRATING	1	2	3	4	5
2. HARD TIME PAYING ATTENTION	1	2	3	4	5
3. CAN'T STAY AT A TASK	1	2	3	4	5
4. DISTRACTIBLE WHILE DOING THINGS	1	2	3	4	5
5. HAVING TO RE-READ MATERIAL	1	2	3	4	5
6. LOSING TRACK OF CONVERSATIONS	1	2	3	4	5
7. HAVING TO PAY EXTRA ATTENTION TO UNDERSTAND	1	2	3	4	5
8. DAYDREAMING	1	2	3	4	5
9. NOT THINKING STRAIGHT/ EFFICIENTLY.	1	2	3	4	5

2. BOREDOM

1. NOTHING SEEMS INTERESTING.	1	2	3	4	5
2. NOT CARING ABOUT WHAT IS HAPPENING.	1	2	3	4	5
3. NONE OF THIS IS FUN ANYMORE.	1	2	3	4	5
4. NOT INTERESTED IN WHAT IS HAPPENING.	1	2	3	4	5
5. INDIFFERENT.	1	2	3	4	5
6. BORED WITH MISSION.	1	2	3	4	5
7. TIRED OF SAME OLD THING.	1	2	3	4	5
8. THINGS ARE TEDIOUS.	1	2	3	4	5
9. TIME IS PASSING TOO SLOWLY.	1	2	3	4	5

3. SLOWED REACTIONS

1. NOT MOVING VERY MUCH.	1	2	3	4	5
2. JUST LOOKING AROUND.	1	2	3	4	5
3. HAVING TO THINK BEFORE ACTING.	1	2	3	4	5
4. THINGS SEEM IN SLOW MOTION.	1	2	3	4	5
5. ARMS FEEL HEAVY.	1	2	3	4	5
6. REACTIONS ARE SLOWED.	1	2	3	4	5
7. MOVEMENTS SEEM DELAYED.	1	2	3	4	5
8. CAN'T KEEP UP WITH TASKS.	1	2	3	4	5
9. RESPONDING TAKES EFFORT.	1	2	3	4	5

4. ANXIETY

1. FEEL ANXIOUS.	1	2	3	4	5
2. FEEL TENSE IN MUSCLES.	1	2	3	4	5
3. FEEL NERVOUS.	1	2	3	4	5
4. WORRIED ABOUT THINGS.	1	2	3	4	5
5. MUSCLES ARE JITTERY.	1	2	3	4	5
6. TAPPING FINGERS OR FOOT.	1	2	3	4	5
7. STOMACH FEELS UPSET.	1	2	3	4	5
8. JUMPY.	1	2	3	4	5
9. VIGILANT.	1	2	3	4	5

5. DEPRESSION

1. FEEL DEPRESSED.	1	2	3	4	5
2. FEEL UNHAPPY.	1	2	3	4	5
3. FEEL SAD.	1	2	3	4	5
4. NOT ENJOYING THIS.	1	2	3	4	5
5. FEEL DISCOURAGED.	1	2	3	4	5
6. WISH I FELT HAPPIER..	1	2	3	4	5
7. WORRIED I MIGHT NEVER FEEL BETTER.	1	2	3	4	5
8. NO CONTROL OVER ANY OF THIS.	1	2	3	4	5
9. WHAT I DO DOESN'T MATTER.	1	2	3	4	5

6. IRRITABILITY

1. FEELING IRRITABLE.	1	2	3	4	5
2. GENERALLY ANNOYED WITH OTHERS.	1	2	3	4	5
3. PRETTY UNFRIENDLY/ TESTY.	1	2	3	4	5
4. IMPATIENT.	1	2	3	4	5
5. WOULD REALLY LIKE TO BE ALONE FOR A WHILE.	1	2	3	4	5
6. GETTING ANGRY.	1	2	3	4	5
7. DISAGREEABLE.	1	2	3	4	5
8. WISH I WEREN'T WITH THESE PEOPLE.	1	2	3	4	5
9. FEELING GROUCHY WITH PEOPLE.	1	2	3	4	5

7. FATIGUE/ LOW ENERGY

1. VERY TIRED.	1	2	3	4	5
2. REALLY FATIGUED.	1	2	3	4	5
3. BURNED OUT.	1	2	3	4	5
4. WORN OUT.	1	2	3	4	5
5. FEELING EXHAUSTED.	1	2	3	4	5
6. CAN'T GO ON MUCH LONGER.	1	2	3	4	5
7. NOT ENERGETIC.	1	2	3	4	5
8. TOO TIRED TO MOVE.	1	2	3	4	5
9. FEEL DRAINED.	1	2	3	4	5

8. POOR SLEEP

1. SLEEPY.	1	2	3	4	5
2. WISH I HAD SLEPT LONGER.	1	2	3	4	5
3. WISH I HAD SLEPT MORE SOUNDLY.	1	2	3	4	5
4. COULD FALL ASLEEP RIGHT HERE.	1	2	3	4	5
5. NODDING OFF.	1	2	3	4	5
6. EYES ARE CLOSING.	1	2	3	4	5
7. FORCING SELF TO KEEP EYES OPEN.	1	2	3	4	5
8. BOUNCING, TAPPING, SHAKING TO STAY AWAKE.	1	2	3	4	5
9. YAWNING.	1	2	3	4	5

9. WORK FRUSTRATION

1. PREFER TO NOT WORK NOW.	1	2	3	4	5
2. WISH I DIDN'T HAVE TO DO THINGS RIGHT NOW.	1	2	3	4	5
3. HOPE NOTHING ELSE NEEDS TO BE DONE.	1	2	3	4	5
4. TOO MUCH IS EXPECTED OF ME RIGHT NOW.	1	2	3	4	5
5. TIRED OF WORKING.	1	2	3	4	5
6. WORK FEELS LIKE DRUDGERY.	1	2	3	4	5
7. FEELS LIKE A GRIND.	1	2	3	4	5
8. WISH THE WORK WOULD END.	1	2	3	4	5
9. WOULD LIKE HELP WITH THE WORK.	1	2	3	4	5

10. PHYSICAL DISCOMFORT

1. MUSCLES HURT.	1	2	3	4	5
2. BODY STIFF.	1	2	3	4	5
3. CAN'T GET COMFORTABLE.	1	2	3	4	5
4. ARMS AND LEGS HURT.	1	2	3	4	5
5. WANT TO STRETCH.	1	2	3	4	5
6. SKIN STICKY/ DIRTY.	1	2	3	4	5
7. WHAT I'M WEARING IS UNCOMFORTABLE..	1	2	3	4	5
8. HEAD ACHES.	1	2	3	4	5
9. EYES TIRED.	1	2	3	4	5

Appendix B

Item-Total Correlations

1. POOR CONCENTRATION	
1. DIFFICULTY CONCENTRATING	.87
2. HARD TIME PAYING ATTENTION	.87
3. CAN'T STAY AT A TASK	.83
4. DISTRACTIBLE WHILE DOING THINGS	.83
5. HAVING TO RE-READ MATERIAL	.80
6. LOSING TRACK OF CONVERSATIONS	.79
7. HAVING TO PAY EXTRA ATTENTION TO UNDERSTAND	.89
8. DAYDREAMING	.68
9. NOT THINKING STRAIGHT/ EFFICIENTLY.	.79
2. BOREDOM	
1. NOTHING SEEMS INTERESTING.	.86
2. NOT CARING ABOUT WHAT IS HAPPENING.	.84
3. NONE OF THIS IS FUN ANYMORE.	.85
4. NOT INTERESTED IN WHAT IS HAPPENING.	.91
5. INDIFFERENT.	.84
6. BORED WITH MISSION.	.85
7. TIRED OF SAME OLD THING.	.82
8. THINGS ARE TEDIOUS.	.82
9. TIME IS PASSING TOO SLOWLY.	.80
3. SLOWED REACTIONS	
1. NOT MOVING VERY MUCH.	.81
2. JUST LOOKING AROUND.	.80
3. HAVING TO THINK BEFORE ACTING.	.61
4. THINGS SEEM IN SLOW MOTION.	.85
5. ARMS FEEL HEAVY.	.83
6. REACTIONS ARE SLOWED.	.88
7. MOVEMENTS SEEM DELAYED.	.86
8. CAN'T KEEP UP WITH TASKS.	.75
9. RESPONDING TAKES EFFORT.	.85

4. ANXIETY

1. FEEL ANXIOUS.	.76
2. FEEL TENSE IN MUSCLES.	.80
3. FEEL NERVOUS.	.87
4. WORRIED ABOUT THINGS.	.78
5. MUSCLES ARE JITTERY.	.82
6. TAPPING FINGERS OR FOOT.	.68
7. STOMACH FEELS UPSET.	.71
8. JUMPY.	.82
9. VIGILANT.	.72

5. DEPRESSION

1. FEEL DEPRESSED.	.87
2. FEEL UNHAPPY.	.85
3. FEEL SAD.	.82
4. NOT ENJOYING THIS.	.81
5. FEEL DISCOURAGED.	.84
6. WISH I FELT HAPPIER..	.79
7. WORRIED I MIGHT NEVER FEEL BETTER.	.68
8. NO CONTROL OVER ANY OF THIS.	.75
9. WHAT I DO DOESN'T MATTER.	.78

6. IRRITABILITY

1. FEELING IRRITABLE.	.83
2. GENERALLY ANNOYED WITH OTHERS.	.89
3. PRETTY UNFRIENDLY/ TESTY.	.82
4. IMPATIENT.	.80
5. WOULD REALLY LIKE TO BE ALONE FOR A WHILE.	.69
6. GETTING ANGRY.	.85
7. DISAGREEABLE.	.81
8. WISH I WEREN'T WITH THESE PEOPLE.	.78
9. FEELING GROUCHY WITH PEOPLE.	.91

7. FATIGUE/ LOW ENERGY

1. VERY TIRED.	.86
2. REALLY FATIGUED.	.89
3. BURNED OUT.	.75
4. WORN OUT.	.89
5. FEELING EXHAUSTED.	.91
6. CAN'T GO ON MUCH LONGER.	.77
7. NOT ENERGETIC.	.84
8. TOO TIRED TO MOVE.	.80
9. FEEL DRAINED.	.88

8. POOR SLEEP

1. SLEEPY.	.81
2. WISH I HAD SLEPT LONGER.	.73
3. WISH I HAD SLEPT MORE SOUNDLY.	.72
4. COULD FALL ASLEEP RIGHT HERE.	.78
5. NODDING OFF.	.82
6. EYES ARE CLOSING.	.77
7. FORCING SELF TO KEEP EYES OPEN.	.78
8. BOUNCING, TAPPING, SHAKING TO STAY AWAKE.	.66
9. YAWNING.	.66

9. WORK FRUSTRATION

1. PREFER TO NOT WORK NOW.	.82
2. WISH I DIDN'T HAVE TO DO THINGS RIGHT NOW.	.91
3. HOPE NOTHING ELSE NEEDS TO BE DONE.	.90
4. TOO MUCH IS EXPECTED OF ME RIGHT NOW.	.85
5. TIRED OF WORKING.	.88
6. WORK FEELS LIKE DRUDGERY.	.91
7. FEELS LIKE A GRIND.	.93
8. WISH THE WORK WOULD END.	.87
9. WOULD LIKE HELP WITH THE WORK.	.73

10. PHYSICAL DISCOMFORT

1. MUSCLES HURT.	.75
2. BODY STIFF.	.78
3. CAN'T GET COMFORTABLE.	.88
4. ARMS AND LEGS HURT.	.76
5. WANT TO STRETCH.	.77
6. SKIN STICKY/ DIRTY.	.54
7. WHAT I'M WEARING IS UNCOMFORTABLE..	.46
8. HEAD ACHES.	.70
9. EYES TIRED.	.76